

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 349 775 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
28.02.1996 Bulletin 1996/09

(51) Int Cl.⁶: **G11C 11/56, G11C 16/04**

(21) Application number: **89110229.5**

(22) Date of filing: **06.06.1989**

(54) **Flash eeprom memory systems and methods of using them**

"Flash"-EEPROM-Speichersysteme und Verfahren zu deren Verwendung

Systèmes de mémoire EEPROM "flash" et méthode les utilisant

(84) Designated Contracting States:
DE FR GB

(30) Priority: **08.06.1988 US 204175**

(43) Date of publication of application:
10.01.1990 Bulletin 1990/02

(60) Divisional application: **95110041.1**

(73) Proprietor: **Harari, Eliyahou**
Los Gatos, CA 95030 (US)

(72) Inventor: **Harari, Eliyahou**
Los Gatos, CA 95030 (US)

(74) Representative: **Blumbach, Kramer & Partner**
D-81245 München (DE)

(56) References cited:
EP-A- 0 251 889 US-A- 4 087 795
US-A- 4 181 980 US-A- 4 279 024
US-A- 4 357 685

- **IEEE JOURNAL OF SOLID-STATE CIRCUITS**, vol. SC-22, no. 3, June 1987, pges 460- 464, IEEE, New York, NY, US; C. BLEIKER et al.: "A four-state EEPROM using floating-gate memory cells"
- **NEW ELECTRONICS**, vol. 15, no. 3, February 1982, pages 46-50, London, GB; P. HAROLD: "Memories; Production E.P.R.O.M. loading"
- **ALTA FREQUENZA**, vol. 52, no. 6, November 1983, pages 487-494, Milano, IT; TORRELLI et al.: "An improved method for programming a word-erasable EEPROM"

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 0 349 775 B1

Description

The invention relates to a method of altering a memory state of an addressed cell of an array according to the preamble of claim 1 and to an array of electrically alterable memory cells according to the preamble of claim 15.

A method of altering according to the preamble of claim 1 and a four-state EEPROM using floating gate memory cells according to the preamble of claim 15 are already known from "IEEE Journal of Solid-State Circuits", Vol. SC-22, No. 3, June 1987, pages 460-464. One memory state is assigned to a positive floating gate voltage and three further states are assigned to negative floating gate voltages. There is no disclosure of a net positive charge on the floating gate being responsive for programmable threshold states of the cell.

In connection with EEPROMs (US-A-4 357 685) it is already known to program the cells by applying charging pulses increasing in height to the floating gate of the cell transistor, and to erase analogously by applying discharging pulses increasing in height. During periods between the single pulses, the cell currents are read to check whether the desired state has been reached. No further condition is disclosed in the prior art.

US-A-4 087 759 discloses a two-state EEPROM having floating gate transistors of the split channel transistor type. The memory states are realized either by positive or by negative charges.

The threshold voltage characteristics of a transistor is dependent from the amount of charge that is retained on the floating gate. That is, the minimum amount of voltage (threshold) that must be applied to the control gate before the transistor is turned "on" to permit conduction between its source and drain regions is determined by the level of charge on the floating gate. A transistor is programmed to one of two or more states by accelerating electrons from the substrate channel region, through a thin gate dielectric and onto the floating gate.

The memory cell transistor's state is read by placing an operating voltage across its source and drain and a selected voltage on its control gate, and then detecting the level of current flowing between the source and drain which determines the programmed state of the transistor.

One example of such a memory cell is a triple polysilicon, split channel electrically erasable and programmable read only memory (EEPROM). It is termed a "split channel" device since the floating and control gates extend over adjacent portions of the channel. This results in a transistor structure that operates as two transistors in series, one having a varying threshold in response to the charge level on the floating gate, and another that is unaffected by the floating gate charge but rather which operates in response to the voltage on the control gate as in any normal field effect transistor.

Such a memory cell is termed a "triple polysilicon" cell because it contains three conductive layers of poly-

silicon materials. In addition to the floating and control gates, an erase gate is included. The erase gate passes through each memory cell transistor closely adjacent to a surface of the floating gate but insulated therefrom by a thin tunnel dielectric. Charge is then removed from the floating gate of a cell to the erase gate, when appropriate voltage are applied to all the transistor elements. An array of EEPROM cells are generally referred to as a Flash EEPROM array if an entire array of cells, or a significant group of cells, is erased simultaneously (i.e. in a flash).

EEPROM's have been found to have a limited effective life. The number of cycles of programming and erasing that such a device can endure before becoming degraded is finite. After a number of such cycles in excess of 10,000, depending upon its specific structure, its programmability can be reduced. Often, by the time the device has been put through such a cycle for over 100,000 times, it can no longer be programmed or erased properly.

This is believed to be the result of electrons being trapped in the dielectric each time charge is transferred to or away from the floating gate by programming or erasing, respectively.

It is a primary object of the present invention to provide an EEPROM array with increased storage capacity and life.

It is a further object of the present invention to provide techniques for increasing the number of program/erase cycles that an EEPROM can endure.

It is an advantage of the present invention to provide EEPROM semiconductor chips that are useful for solid state memory to replace magnetic disk storage devices.

The features of the invention are specified in claims 1 and 15.

The present invention relates, briefly and generally, to an EEPROM memory cell which is caused to store more than one bit of data by partitioning its programmed charge into three or more ranges. Each cell is then programmed into one of these ranges. If four ranges are used, two bits of data can be stored in a single cell. If eight ranges are designated, three bits can be stored, and so on.

An intelligent programming and sensing technique is provided which permits the practical implementation of such multiple state storage. Further, an intelligent erase algorithm is provided which results in a significant reduction in the electrical stress experienced by the erase tunnel dielectric and results in much higher endurance to program/erase cycling and a resulting increased life of the memory.

Additional objects, features and advantages of the present invention will be understood from the following description of its preferred embodiments, which description should be taken in conjunction with the accompanying drawings.

Brief Description of the Drawings

Figure 1 is a cross section of an example split channel Eprom or EEPROM.

Figure 2a is a schematic representation of the composite transistor forming a split channel Eprom device.

Figure 2b shows the programming and erase characteristics of a split channel Flash EEPROM device.

Figure 2c shows the four conduction states of a split channel Flash EEPROM device in accordance with this invention.

Figure 2d shows the program/erase cycling endurance characteristics of prior art Flash EEPROM devices.

Figure 2e shows a circuit schematic and programming/read voltage pulses required to implement multi-state storage.

Figure 3 outlines the key steps in the new algorithm used to erase with a minimum stress.

Figure 4 shows the program/erase cycling endurance characteristics of the split channel Flash EEPROM device using intelligent algorithms for multistate programming and for reduced stress during erasing.

Description of the Preferred Embodiments

Referring initially to Figure 1, the structure of a split-channel Eprom or EEPROM cell is described that is suitable for use in the improved memory array and operation of the present invention. A semiconductor substrate 11 includes source region 13 and drain region 15, usually formed by ion implantation. Between the source and drain is a channel region 17. Over a portion of the channel region designated as L1 is a floating gate 19, separated from the substrate by a thin layer of gate oxide 21. Over a portion of the channel region designated as L2 is formed a control gate 23, separated from the substrate 11 by a thin gate oxide layer 25. The control gate 23 is also electrically isolated from the floating gate 19 by an oxide layer 27.

It is the amount of electrical charge on the floating gate 19 that is programmed in accordance with the state desired to be stored in the cell. If the charge level is above some set threshold, the cell is considered to be in one state. If below that threshold, it is designated to be in its other state. The desired charge level is programmed by applying an appropriate combination of voltages to the source, drain, substrate and control gate, for a designated period of time, in order to cause electrons to move from the substrate 11 to the floating gate 19.

The floating gate is confined to its one memory cell and is electrically isolated from all other parts of the structure. The control gate 23, on the other hand, extends across a large number of cells sharing a common word line. As described hereinafter, the split-channel has the effect of providing two field-effect-transistors in series, one with the floating gate 19 as its control gate and the other with the control gate 23 as its control gate.

The generic split-channel Eprom or EEPROM structure of Figure 1 becomes a Flash EEPROM device when an erase gate (not shown) is added. The erase gate is a separate electrode positioned near a portion of the floating gate 27 and separated from it by a tunnel dielectric. When the proper voltages are applied to the source, drain, substrate, control gate and erase gate, the amount of charge on the floating gate is reduced. A single erase gate extends to a large number of memory cells, if not the entire array, so that they may be erased all at once. In some prior art Flash EEPROM cells the source or drain diffusions underneath the floating gate are used also as the erase electrode, while in other cells the erase electrode is implemented either in the same conductive layer as the control gate or in a separate conductive layer.

Multistate storage

The split channel Flash EEPROM device can be viewed as a composite transistor consisting of two transistors T1 and T2 in series - figure 2a. Transistor T1 is a floating gate transistor of effective channel length L1 and having a variable threshold voltage V_{T1} . Transistor T2 has a fixed (enhancement) threshold voltage V_{T2} and an effective channel length L2. The Eprom programming characteristics of the composite transistor are shown in curve (a) of figure 2b. The programmed threshold voltage V_{tx} is plotted as a function of the time t during which the programming conditions are applied. These programming conditions typically are $V_{CG} = 12V$, $V_D = 9V$, $V_S = V_{BB} = 0V$. No programming can occur if either one of V_{CG} or V_D is at 0V. A Virgin (unprogrammed, un-erased) device has $V_{T1} = +1.5V$ and $V_{T2} = +1.0V$. After programming for approximately 100 microseconds the device reaches a threshold voltage $V_{tx} \geq +6.0$ volts. This represents the off ("0") state because the composite device does not conduct at $V_{CG} = +5.0V$. Prior art devices employ a so called "intelligent programming" algorithm whereby programming pulses are applied, each of typically 100 microseconds to 1 millisecond duration, followed by a sensing (read) operation. Pulses are applied until the device is sensed to be fully in the off state, and then one to three more programming pulses are applied to ensure solid programmability.

Prior art split channel Flash EEPROM devices erase with a single pulse of sufficient voltage V_{ERASE} and sufficient duration to ensure that V_{T1} is erased to a voltage below V_{T2} (curve (b) in figure 2b). Although the floating gate transistor may continue to erase into depletion mode operation (line (c) in figure 2b), the presence of the series T2 transistor obscures this depletion threshold voltage. Therefore the erased on ("1") state is represented by the threshold voltage $V_{tx} = V_{T2} = +1.0V$. The memory storage "window" is given by $\Delta V = V_{tx}("0") - V_{tx}("1") = 6.0 - 1.0 = 5.0V$. However, the true memory storage window should be represented by the full swing of V_{tx} for transistor T1. For example, if T1 is erased into depletion threshold voltage $V_{T1} = -3.0V$, then the true window

should be given by $\Delta V = 6.0 - (-3.0) = 9.0V$. None of the prior art Flash EEPROM devices take advantage of the true memory window. In fact they ignore altogether the region of device operation (hatched region D in figure 2b) where V_{T1} is more negative than V_{T2} .

This invention proposes for the first time a scheme to take advantage of the full memory window. This is done by using the wider memory window to store more than two binary states and therefore more than a single bit per cell. For example, it is possible to store 4, rather than 2 states per cell, with these states having the following threshold voltage:

State "3": - $V_{T1} = -3.0V$, $V_{T2} = +1.0V$
(highest conduction) = 1, 1.

State "2": - $V_{T1} = -0.5V$, $V_{T2} = +1.0V$
(intermediate conduction) = 1, 0.

State "1": - $V_{T1} = +2.0V$, $V_{T2} = +1.0V$
(lower conduction) = 0, 1.

State "0": - $V_{T1} = +4.5V$, $V_{T2} = +1.0V$
(no conduction) = 0, 0.

To sense any one of these four states, the control gate is raised to $V_{CG} = +5.0V$ and the source-drain current I_{DS} is sensed through the composite device. Since $V_{T2} = +1.0V$ for all four threshold states transistor T2 behaves simply as a series resistor. The conduction current I_{DS} of the composite transistor for all 4 states is shown as a function of V_{CG} in figure 2c. A current sensing amplifier is capable of easily distinguishing between these four conduction states. The maximum number of states which is realistically feasible is influenced by the noise sensitivity of the sense amplifier as well as by any charge loss which can be expected over time at elevated temperatures. Eight distinct conduction states are necessary for 3 bit storage per cell, and 16 distinct conduction states are required for 4 bit storage per cell.

Multistate memory cells have previously been proposed in conjunction with ROM (Read Only Memory) devices and DRAM (Dynamic Random Access Memory). In ROM, each storage transistor can have one of several fixed conduction states by having different channel ion implant doses to establish more than two permanent threshold voltage states. Prior art multistate DRAM cells have also been proposed where each cell in the array is physically identical to all other cells. However, the charge stored at the capacitor of each cell may be quantized, resulting in several distinct read signal levels. An example of such prior art multistate DRAM storage is described in IEEE Journal of Solid-State Circuits, Feb. 1988, p. 27 in an article by M. Horiguchi et al. entitled "An Experimental Large-Capacity Semiconductor File Memory Using 16-Levels/Cell Storage". A second example of prior art multistate DRAM is provided in IEEE Custom Integrated Circuits Conference, May 1988, p. 4.4.1

in an article entitled "An Experimental 2-Bit/Cell Storage DRAM for Macrocell or Memory-on-Logic Applications" by T. Furuyama et al.

To take full advantage of multistate storage in Eeproms it is necessary that the programming algorithm allow programming of the device into any one of several conduction states. First it is required that the device be erased to a voltage V_{T1} more negative than the "3" state ($-3.0V$ in this example). Then the device is programmed in a short programming pulse, typically one to ten microseconds in duration. Programming conditions are selected such that no single pulse can shift the device threshold by more than one half of the threshold voltage difference between two successive states. The device is then sensed by comparing its conduction current I_{DS} with that of a reference current source I_{REF} , i ($i=0,1,2,3$) corresponding to the desired conduction state (four distinct reference levels must be provided corresponding to the four states). Programming pulses are continued until the sensed current (solid lines in figure 2c) drops slightly below the reference current corresponding to the desired one of four states (dashed lines in figure 2c). To better illustrate this point, assume that each programming pulse raises V_{tx} linearly by 200 millivolts, and assume further that the device was first erased to $V_{T1} = -3.2V$. Then the number of programming/sensing pulses required is:

For state "3" ($V_{T1} = -3.0V$)
No. of pulses = $(3.2-3.0)/.2 = 1$

For state "2" ($V_{T1} = -0.5V$)
No. of pulses = $(3.2-0.5)/.2 = 14$

For state "1" ($V_{T1} = +2.0V$)
No. of pulses = $(3.2-(-2.0))/.2 = 26$

and for state "0" ($V_{T1} = +4.5V$)
No. of pulses = $(3.2-(-4.5))/.2 = 39$.

In actual fact shifts in V_{tx} are not linear in time, as shown in figure 2b (curve (a)), therefore requiring more pulses than indicated for states "1" and "0". If 2 microseconds is the programming pulse width and 0.1 microseconds is the time required for sensing, then the maximum time required to program the device into any of the 4 states is approximately $39 \times 2 + 39 \times .1 = 81.9$ microseconds. This is less than the time required by "intelligent programming algorithms" of prior art devices. In fact, with the new programming algorithm only carefully metered packets of electrons are injected during programming. A further benefit of this approach is that the sensing during reading is the same sensing as that during programming/sensing, and the same reference current sources are used in both programming and reading operations. That means that each and every memory cell in the array is read relative to the same reference level as used during program/sense. This provides excellent tracking even in

very large memory arrays.

Large memory systems typically incorporate error detection and correction schemes which can tolerate a small number of hard failures i.e. bad Flash EEPROM cells. For this reason the programming/sensing cycling algorithm can be automatically halted after a certain maximum number of programming cycles has been applied even if the cell being programmed has not reached the desired threshold voltage state, indicating a faulty memory cell.

There are several ways to implement the multistate storage concept in conjunction with an array of Flash EEPROM transistors. An example of one such circuit is shown in figure 2c. In this circuit an array of memory cells has decoded word lines and decoded bit lines connected to the control gates and drains respectively of rows and columns of cells. Each bit line is normally precharged to a voltage of between 1.0 V and 2.0 V during the time between read, program or erase. For a four state storage, four sense amplifiers, each with its own distinct current reference levels $I_{REF,0}$, $I_{REF,1}$, $I_{REF,2}$, and $I_{REF,3}$ are attached to each decoded output of the bit line. During read, the current through the Flash EEPROM transistor is compared simultaneously (i.e., in parallel) with these four reference levels (this operation can also be performed in four consecutive read cycles using a single sense amplifier with a different reference applied at each cycle, if the attendant additional time required for reading is not a concern). The data output is provided from the four sense amplifiers through four Di buffers (D0, D1, D2 and D3).

During programming, the four data inputs I_i (I0, I1, I2 and I3) are presented to a comparator circuit which also has presented to it the four sense amp outputs for the accessed cell. If D_i match I_i , then the cell is in the correct state and no programming is required. If however all four D_i do not match all four I_i , then the comparator output activates a programming control circuit. This circuit in turn controls the bit line (VPBL) and word line (VPWL) programming pulse generators. A single short programming pulse is applied to both the selected word line and the selected bit line. This is followed by a second read cycle to determine if a match between D_i and I_i has been established. This sequence is repeated through multiple programming/reading pulses and is stopped only when a match is established (or earlier if no match has been established but after a preset maximum number of pulses has been reached).

The result of such multistate programming algorithm is that each cell is programmed into any one of the four conduction states in direct correlation with the reference conduction states $I_{REF,i}$. In fact, the same sense amplifiers used during programming/reading pulsing are also used during sensing (i.e., during normal reading). This allows excellent tracking between the reference levels (dashed lines in figure 2c) and the programmed conduction levels (solid lines in figure 2c), across large memory arrays and also for a very wide range of operating

temperatures. Furthermore, because only a carefully metered number of electrons is introduced onto the floating gate during programming or removed during erasing, the device experiences the minimum amount of endurance-related stress possible.

In actual fact, although four reference levels and four sense amplifiers are used to program the cell into one of four distinct conduction states, only three sense amplifiers and three reference levels are required to sense the correct one of four stored states. For example, in figure 2c, $I_{REF}("2")$ can differentiate correctly between conduction states "3" and "2", $I_{REF}("1")$ can differentiate correctly between conduction states "2" and "1", and $I_{REF}("0")$ can differentiate correctly between conduction states "1" and "0". In a practical implementation of the circuit of figure 2e the reference levels $I_{REF,i}$ ($i=0,1,2$) may be somewhat shifted by a fixed amount during sensing to place them closer to the midpoint between the corresponding lower and higher conduction states of the cell being sensed.

Note that the same principle employed in the circuit of figure 2e can be used also with binary storage, or with storage of more than four states per cell. Of course, circuits other than the one shown in figure 2e are also possible. For example, voltage level sensing rather than conduction level sensing can be employed.

Improved Charge Retention

In the example above, states "3" and "2" are the result of net positive charge (holes) on the floating gate while states "1" and "0" are the result of net negative charge (electrons) on the floating gate. To properly sense the correct conduction state during the lifetime of the device (which may be specified as 10 years at 125°C) it is necessary for this charge not to leak off the floating gate by more than the equivalent of approximately 200 millivolts shift in V_{T1} . This condition is readily met for stored electrons in this as well as all prior art EPROM and Flash EEPROM devices. From device physics considerations, retention of holes trapped on the floating gate should be significantly superior to the retention of trapped electrons. This is so because trapped holes can only be neutralized by the injection of electrons onto the floating gate. So long as the conditions for such injection do not exist it is almost impossible for the holes to overcome the potential barrier of approximately 5.0 electronvolts at the silicon-silicon dioxide interface (compared to a 3.1 electron volts potential barrier for trapped electrons).

Therefore it is possible to improve the retention of this device by assigning more of the conduction states to states which involve trapped holes. For example, in the example above state "1" had $V_{T1} = +2.0V$, which involved trapped electrons since V_{T1} for the virgin device was made to be $V_{T1} = +1.5V$. If however V_{T1} of the virgin device is raised to a higher threshold voltage, say to $V_{T1} = +3.0V$ (e.g. by increasing the p-type doping concentration in the channel region 560a in figure 5a), then the

same state "1" with $V_{T1} = +2.0V$ will involve trapped holes, and will therefore better retain this value of V_{T1} . Of course it is also possible to set the reference levels so that most or all states will have values of V_{T1} which are lower than the V_{T1} of the virgin device.

Intelligent Erase for Improved Endurance

The endurance of Flash EEPROM devices is their ability to withstand a given number of program/erase cycles. The physical phenomenon limiting the endurance of prior art Flash EEPROM devices is trapping of electrons in the active dielectric films of the device. During programming the dielectric used during hot electron channel injection traps part of the injected electrons. During erasing the tunnel erase dielectric likewise traps some of the tunneled electrons. The trapped electrons oppose the applied electric field in subsequent write/erase cycles thereby causing a reduction in the threshold voltage shift of V_{tx} . This can be seen in a gradual closure (figure 2d) in the voltage "window" between the "0" and "1" states. Beyond approximately 1×10^4 program/erase cycles the window closure can become sufficiently severe to cause the sensing circuitry to malfunction. If cycling is continued the device eventually experiences catastrophic failure due to a ruptured dielectric. This typically occurs at between 1×10^6 and 1×10^7 cycles, and is known as the intrinsic breakdown of the device. In memory arrays of prior art devices the window closure is what limits the practical endurance to approximately 1×10^4 cycles. At a given erase voltage, V_{ERASE} , the time required to adequately erase the device can stretch out from 100 milliseconds initially (i.e. in a virgin device) to 10 seconds in a device which has been cycled through 1×10^4 cycles. In anticipation of such degradation prior art Flash EEPROM devices specified to withstand 1×10^4 cycles must specify a sufficiently long erase pulse duration to allow proper erase after 1×10^4 cycles. However this also results in virgin devices being overerased and therefore being unnecessarily overstressed.

A second problem with prior art devices is that during the erase pulse the tunnel dielectric may be exposed to an excessively high peak stress. This occurs in a device which has previously been programmed to state "0" ($V_{T1} = +4.5V$ or higher). This device has a large negative charge Q . When V_{ERASE} is applied the tunnel dielectric is momentarily exposed to a peak electric field with contributions from V_{ERASE} as well as from Q . This peak field is eventually reduced when Q is reduced to zero as a consequence of the tunnel erase. Nevertheless, permanent and cumulative damage is inflicted through this erase procedure, which brings about premature device failure.

To overcome the two problems of overstress and window closure a new erase algorithm is disclosed, which can also be applied equally well to any prior art Flash EEPROM device. Without such new erase algorithm it would be difficult to have a multistate device since, from

curve (b) in figure 2d, conduction states having V_{T1} more negative than V_{T2} may be eliminated after 1×10^4 to 1×10^5 write/erase cycles.

Figure 3 outlines the main steps in the sequence of the new erase algorithm. Assume that a block array of $m \times n$ memory cells is to be fully erased (Flash erase) to state "3" (highest conductivity and lowest V_{T1} state). Certain parameters are established in conjunction with the erase algorithm. They are listed in figure 3: V_1 is the erase voltage of the first erase pulse. V_1 is lower by perhaps 5 volts from the erase voltage required to erase a virgin device to state "3" in a one second erase pulse. t is chosen to be approximately 1/10th of the time required to fully erase a virgin device to state "3". Typically, V_1 may be between 10 and 20 volts while t may be between 10 and 100 milliseconds. The algorithm assumes that a certain small number, X , of bad bits can be tolerated by the system (through for example error detection and correction schemes implemented at the system level. If no error detection and correction is implemented then $X = 0$). These would be bits which may have a shorted or leaky tunnel dielectric which prevents them from being erased even after a very long erase pulse. To avoid excessive erasing the total number of erase pulses in a complete block erase cycling can be limited to a preset number, n_{max} . ΔV is the voltage by which each successive erase pulse is incremented. Typically, ΔV is in the range between 0.25V and 1.0V. For example, if $V_1 = 15.0V$ and $\Delta V = 1.0V$, then the seventh erase pulse will be of magnitude $V_{ERASE} = 21.0V$ and duration t . A cell is considered to be fully erased when its read conductance is greater than I_{30} . The number S of complete erase cyclings experienced by each block is an important information at the system level. If S is known for each block then a block can be replaced automatically with a new redundant block once S reaches 1×10^6 (or any other set number) of program/erase cycles. S is set at zero initially, and is incremented by one for each complete block erase multiple pulse cycle. The value of S at any one time can be stored by using for example twenty bits (2^{20} equals approximately 1×10^6) in each block. That way each block carries its own endurance history. Alternatively the S value can be stored off chip as part of the system.

The sequence for a complete erase cycle of the new algorithm is as follows (see figure 3):

1. Read S . This value can be stored in a register file. (This step can be omitted if S is not expected to approach the endurance limit during the operating lifetime of the device).
- 1a. Apply a first erase pulse with $V_{ERASE} = V_1 + n \Delta V$, $n=0$, pulse duration = t . This pulse (and the next few successive pulses) is insufficient to fully erase all memory cells, but it serves to reduce the charge Q on programmed cells at a relatively low erase field stress, i.e., it is equivalent to a "conditioning" pulse.
- 1b. Read a sparse pattern of cells in the array. A

diagonal read pattern for example will read $m+n$ cells (rather than $m \times n$ cells for a complete read) and will have at least one cell from each row and one cell from each column in the array. The number N of cells not fully erased to state "3" is counted and compared with X .

1c. If N is greater than x (array not adequately erased) a second erase pulse is applied of magnitude greater by ΔV than the magnitude of the first pulse, with the same pulse duration, t . Read diagonal cells, count N .

This cycling of erase pulse/read/increment erase pulse is continued until either $N \leq X$ or the number n of erase pulses exceed n_{\max} . The first one of these two conditions to occur leads to a final erase pulse.

2a. The final erase pulse is applied to assure that the array is solidly and fully erased. The magnitude of V_{ERASE} can be the same as in the previous pulse or higher by another increment ΔV . The duration can be between $1t$ and $5t$.

2b. 100% of the array is read. The number N of cells not fully erased is counted. If N is less than or equal to X , then the erase pulsing is completed at this point.

2c. If N is greater than X , then address locations of the N unerased bits are generated, possibly for substitution with redundant good bits at the system level. If N is significantly larger than X (for example, if N represents perhaps 5% of the total number of cells), then a flag may be raised, to indicate to the user that the array may have reached its endurance end of life.

2d. Erase pulsing is ended.

3a. S is incremented by one and the new S is stored for future reference. This step is optional. The new S can be stored either by writing it into the newly erased block or off chip in a separate register file.

3b. The erase cycle is ended. The complete cycle is expected to be completed with between 10 to 20 erase pulses and to last a total of approximately one second.

The new algorithm has the following advantages:

(a) No cell in the array experiences the peak electric field stress. By the time V_{ERASE} is incremented to a relatively high voltage any charge Q on the floating gates has already been removed in previous lower voltage erase pulses.

(b) The total erase time is significantly shorter than the fixed V_{ERASE} pulse of the prior art. Virgin devices see the minimum pulse duration necessary to erase. Devices which have undergone more than 1×10^4 cycles require only several more ΔV voltage increments to overcome dielectric trapped charge, which only adds several hundred milliseconds to their total

erase time.

(c) The window closure on the erase side (curve (b) in figure 2d) is avoided indefinitely (until the device experiences failure by a catastrophic breakdown) because V_{ERASE} is simply incremented until the device is erased properly to state "3". Thus, the new erase algorithm preserves the full memory window.

Figure 4 shows the four conduction states of the Flash EEPROM devices of this invention as a function of the number of program/erase cycles. Since all four states are always accomplished by programming or erasing to fixed reference conduction states, there is no window closure for any of these states at least until 1×10^6 cycles.

In a Flash EEPROM memory chip it is possible to implement efficiently the new erase algorithm by providing on chip (or alternatively on a separate controller chip) a voltage multiplier to provide the necessary voltage V_1 and voltage increments ΔV to $n\Delta V$, timing circuitry to time the erase and sense pulse duration, counting circuitry to count N and compare it with the stored value for X , registers to store address locations of bad bits, and control and sequencing circuitry, including the instruction set to execute the erase sequence outlined above.

While the embodiments of this invention that have been described are the preferred implementations, those skilled in the art will understand that variations thereof may also be possible. Therefore, the invention is entitled to protection within the full scope of the appended claims.

Claims

1. A method of altering a memory state of an addressed cell of an array, the array of electrically alterable memory cells having means for addressing individual memory cells to read and alter their states, each cell having a field effect transistor with a floating gate (19) and having a threshold voltage level that is a given level in the absence of net charge on said floating gate, but which is variable in accordance with an amount of net charge carried by said floating gate, comprising the steps of:

establishing a plurality of effective threshold voltage levels (V_{T1}) in excess of two that correspond to a plurality of individually detectable states ("0", "1", "2", "3") of the cell in excess of two,

presetting the effective threshold voltage of the addressed cell to a predetermined level by altering the amount of charge on the floating gate, and

setting the addressed cell to one of its said plurality of states by altering the amount of charge on the floating gate until its effective threshold voltage is substantially equal to one of said plu-

rality of effective threshold voltage levels, characterized in that a majority of said plurality of effective threshold levels or at least two effective threshold levels ("2", "3") result from a net positive charge on the floating gate.

2. The method of claim 1 wherein the step of establishing a plurality of effective threshold voltage levels V_{T1} includes establishing at least four such effective threshold voltage levels.
3. The method of claim 1 or 2 wherein the threshold voltage level establishing step includes selecting all of said plurality of effective threshold voltage levels to result from a net positive charge on the floating gate.
4. The method of any of claims 1 through 3 wherein the given threshold level of the memory cell transistors is at least three volts.
5. The method of any of claims 1 through 4 which comprises the additional step of accumulating a count (S) equal to a total number of times that the cell has been preset.
6. The method of any of claims 1 through 5 wherein the presetting step is an erasing step carried out by removing negative charge from the floating gate (19) so as to lower the effective threshold voltage of the cell to a base level that is lower than the lowest level corresponding to said plurality of detectable memory states of the cell, and the setting step is a programming step carried out by adding negative charge to the floating gate (19).
7. The method of claim 6 wherein the step of programming the cell to any one of its said plurality of threshold voltage levels includes alternately pulsing the cell with short programming pulse and then reading the current through the cell, continuing with pulsing and reading until the current level reaches that corresponding to the desired one of said plurality of effective threshold levels, each one of said short programming pulses being insufficient to change said effective threshold voltage by more than approximately one-half of the difference between any two adjacent of said plurality of threshold voltage levels.
8. The method of claim 6 wherein the step of erasing the cell includes alternately pulsing the cell with an erase pulse and reading the current through the cell, continuing with pulsing and reading until the current level reaches that corresponding to the desired base threshold level, the magnitude and duration of each such erase

pulse being chosen so that the first erase pulse is insufficient to fully erase said cell and each subsequent erase pulse has its magnitude increased by a fixed increment until the cell is fully erased.

9. The method of claim 7 wherein the step of erasing the cell includes alternately pulsing the cell with an erase pulse and reading its effective threshold level, continuing with pulsing and reading until either said effective threshold level reaches the desired base threshold level or the number of erase pulses applied has exceeded a preset number.
10. The method according to claim 5 wherein said cells of said array form blocks of cells to be addressed and further comprising the additional step of storing said count (S) in count storing cells within each block of cells to which the count pertains.
11. The method according to claim 10 wherein said count (S) is temporarily stored in a register file, then the addressed block of cells is erased and the count (S) is incremented by one, whereupon the incremented count is rewritten into said count storing cells of the addressed block.
12. The method of claim 6 and 10 wherein the steps of erasing an addressed block includes:
 - pulsing the cells for a predetermined time and voltage sufficient to alter the threshold voltage but insufficient to completely erase said cells, thereafter reading the current through a selected number (m + n) of cells in order to ascertain their altered threshold voltage, and repeating the pulsing and reading steps a plurality of times, each repeat of the pulsing step increasing the predetermined voltage an increment above that of the last pulsing step, terminating the pulsing and reading steps upon the first occurrence of any one of the following conditions: the effective threshold voltage of each of said selected number (m + n) of cells has reached the base level; or the pulsing step has been repeated a preset maximum number of times; or a predetermined maximum voltage for a pulse has been reached in the most recent pulsing step; or a number (N) of cells of said selected number (m + n) of cells which remain not fully erased is equal to or less than an acceptable number of unerased cells.
13. The method according to claim 12

wherein the selected number ($m + n$) of cells is significantly less than the total number ($m \times n$) of the cells of the addressed block.

14. The method according to claim 12 or 13 wherein the count (S) equal to the total number of erase cycles that the cells of the addressed block have experienced, is accumulated and stored in count storing cells within each block of cells to which the count (S) pertains.

15. A device in the form of an array of electrically alterable memory cells having means for addressing individual memory cells to read and alter their states, each cell including a field effect transistor with

a source region (13),
a drain region (15),
a channel region (17),
a floating gate (19), and
a control gate (23),
the field effect transistor having a threshold voltage level that is a given level in the absence of net charge on said floating gate but which is variable in accordance with an amount of net charge carried by said floating gate,
the device comprising
means for establishing a plurality of effective threshold voltage levels (V_{T1}) in excess of two that correspond to a plurality of individually detectable states ("0", "1", "2", "3") of the cell in excess of two,
means for presetting the effective threshold voltage of the addressed cell to a predetermined level by altering the amount of charge on the floating gate, and
means for setting the addressed cell to one of its said plurality of states by altering the amount of charge on the floating gate until its effective threshold voltage is substantially equal to one of said plurality of effective threshold voltage levels,
characterized in that a majority of said plurality of effective threshold levels or at least two effective threshold levels ("2", "3") result from a net positive charge on the floating gate (19).

16. The device of claim 15

wherein said field effect transistor is of the split channel type with a first channel portion (L_1) and a second channel portion (L_2), and
wherein said floating gate (19) is positioned adjacent said drain region (15) and extends across the first portion (L_1) of said channel region (17) with a first gate dielectric layer (21) therebetween in a manner to control the amount of conduction through the first channel portion

(L_1) in accordance with a level of electron charge stored on said floating gate (19),
said control gate (23) extends over at least a portion of the floating gate (19) and being insulated (27) therefrom, thereby forming a first transistor (T_1) having said threshold voltage level (V_{T1}) that is given in the absence of net charge on said floating gate (19) but which is variable in accordance with an amount of net charge carried by said floating gate (19),
said control gate (23) is also positioned adjacent said source region (13) and extends across the second portion (L_2) of said channel region (17) with a second gate dielectric layer (25) therebetween in a manner to control the conduction through the second channel portion (L_2) in accordance with a level of voltage applied to the control gate (23), thereby forming a second transistor (T_2) having a fixed threshold voltage (V_{T2}), and
wherein said threshold level (V_{T1}) given by said first transistor (T_1) is established sufficiently high so that at least two programmable threshold levels result from a net positive charge on the floating gate (19).

17. The device according to claim 16 wherein said given threshold level (V_{T1}) of said first transistor (T_1) is greater than the fixed threshold (V_{T2}) of said second transistor (T_2).
18. The device according to claims 16 or 17 wherein said given threshold level (V_{T1}) of said first transistor (T_1) is at least 3 volts.
19. The device according to claim 18 wherein the fixed threshold level (V_{T2}) of the second transistor portion (T_2) is approximately 1 volt.
20. The device according to any of claims 16 through 17 wherein said given threshold level (V_{T1}) of said first transistor (T_1) is established sufficiently high in order that the majority or all of said programmable threshold levels result from a net positive charge on the floating gate (19).
21. The device according to any of claims 16 through 20 wherein said plurality of programmable threshold levels of the first transistor (T_1) is exactly three.
22. The device according to any one of claims 16 to 20 wherein said plurality of programmable threshold levels of the first transistor (T_1) is exactly four.
23. The device according to any one of claims 15 to 22 wherein said presetting and setting means comprise:

erasing means electrically connected with at least one erase electrode of an addressed cell for causing electron charge to be removed from the floating gate (19) until the floating gate (19) becomes positively charged to a maximum amount,

programming means electrically connected with at least said source (13), drain (15) and control gate (23) for causing electron charge to be added to the floating gate (19) to bring the amount of electronic charge stored thereon from said maximum positive charge to a desired one of said effective threshold voltage levels, and

reading means electrically connected with at least said source (13), drain (15) and control gate (23) for reading the level of charge stored on the floating gate (19) and determining in which of said effective threshold voltage levels it lies, thereby to read the individual detectable state ("0", "1", "2", "3") of the cell.

24. The device according of claim 23

wherein said erasing means are simultaneously connectable to erase electrodes or gates of storage cells within one of a plurality of distinct blocks of cells in multiple rows and columns of the array for simultaneously altering the electron charge from the floating gates (19) of cells within said one block,

wherein said erasing means applies a sequence of pulses of increasing magnitude to said erase electrodes or gates of said one block until at least either

(i) the threshold levels (VT_1) detected by said reading means between said erase pulses existing in at least a preset proportion ($N \leq X$) of the cells of said one block reach an erased level, or

(ii) a preset maximum number (n_{max}) of erase pulses have occurred, and

wherein said programming means is connectable to an addressed cell within the array of cells for increasing the electron charge on its floating gate (19) by applying a sequence of programming pulses to the addressed cell until at least either

(iii) the threshold level detected by said reading means between said programming pulses existing in the addressed cell is raised from said erase level to substantially equal a desired one of said plurality of predetermined threshold levels (V_{T1}), or

(iv) a preset maximum number of program-

ming pulses has been applied.

25. The device of claim 24

wherein said programming means includes means for limiting a magnitude of each of said sequence of programming pulses to a value that is less than that which shifts the threshold level of the first transistor (T_1) of claim 16 one-half a difference between adjacent ones of said plurality of effective threshold voltage levels.

26. The device of any of claims 15 to 25

further comprising reading means electrically connected with at least said source (13), drain (15) and control gate (23) for reading the level of charge stored on the floating gate (19), wherein said reading means includes a plurality of reference current sources corresponding respectively to said plurality of effective threshold voltage levels (VT_1), and means for simultaneously comparing current flowing through the addressed cell with each of the reference current sources, thereby to determine the memory state ("0", "1", "2", "3") of the addressed cell.

Patentansprüche

1. Verfahren zur Änderung eines Speicherzustandes einer adressierten Zelle einer Anordnung, wobei die Anordnung der elektrisch änderbaren Speicherzellen Einrichtungen zur Adressierung individueller Speicherzellen aufweist, um die Zustände zu lesen und zu ändern, wobei jede Zelle einen Feldeffekttransistor mit einem schwimmenden Gate (19) aufweist und einen Schwellwertspannungspegel besitzt, der bei Abwesenheit einer Nettoladung auf dem schwimmenden Gate auf einem gegebenen Pegel ist, der jedoch gemäß einem Betrag von Nettoladung variabel ist, die von dem schwimmenden Gate geführt wird, mit folgenden Schritten:

eine Mehrzahl von effektiven Schwellwertspannungspegeln (V_{T1}) oberhalb von Zwei wird errichtet, die einer Mehrzahl von individuell feststellbaren Zuständen ("0", "1", "2", "3") der Zelle oberhalb von Zwei entspricht, der effektive Spannungspegel der adressierten Zelle wird auf einen vorbestimmten Pegel dadurch voreingestellt, daß der Ladungsbetrag auf dem schwimmenden Gate geändert wird,

die adressierte Zelle wird auf einen Zustand der Mehrzahl möglicher Zustände dadurch gesetzt, daß der Ladungsbetrag auf dem schwimmenden Gate geändert wird, bis die effektive

- Schwellwertspannung im wesentlichen gleich einem dieser effektiven Schwellwertspannungspegel der Mehrzahl möglicher Pegel ist, dadurch gekennzeichnet, daß eine Majorität der Mehrzahl der effektiven Schwellwertpegel oder mindestens zwei effektive Schwellwertpegel ("2", "3") von einer nettopositiven Ladung auf dem schwimmenden Gate resultiert.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß der Schritt der Errichtung einer Mehrzahl von effektiven Schwellwertspannungspegeln V_{T1} beinhaltet, daß mindestens vier solcher effektiven Schwellwertspannungspegel errichtet werden.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß gemäß dem Schritt der Errichtung des Schwellwertspannungspegels alle effektiven Schwellwertspannungspegel der Mehrzahl möglicher Pegel so gewählt werden, daß sie von einer nettopositiven Ladung auf dem schwimmenden Gate resultieren.
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß der gegebene Schwellwertpegel der Transistoren der Speicherzelle mindestens 3 V beträgt.
5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß der zusätzliche Schritt der Ansammlung eines Zählstandes (S) gleich der Gesamtanzahl der Voreinstellungen der Zelle vorgesehen ist.
6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet,
- daß der Voreinstellschritt ein Löschungsschritt ist, der durch Abfuhr negativer Ladung von dem schwimmenden Gate (19) ausgeführt wird, um so die effektive Schwellwertspannung der Zelle auf einen Basispegel abzusinken, der niedriger ist als der niedrigste Pegel entsprechend der Mehrzahl der feststellbaren Speicherzustände der Zelle, und
- daß der Einstellungsschritt einen Programmierschritt darstellt, der durch Zufuhr negativer Ladung an das schwimmende Gate (19) ausgeführt wird.
7. Verfahren nach Anspruch 6, dadurch gekennzeichnet,
- daß gemäß dem Schritt der Programmierung der Zelle auf einer ihrer Schwellwertspannungspegel die Zelle mit kurzen Programmierimpulsen abwechselnd gepulst und dann der Strom durch die Zelle gelesen wird, daß mit Pulsen und Lesen fortgefahren wird, bis der Strompegel den Pegel entsprechend dem einen gewünschten effektiven Schwellwertpegel der Mehrzahl möglicher Pegel erreicht, wobei jeder der kurzen Programmierimpulse ungenügend ist, um die effektive Schwellwertspannung um mehr als ungefähr die Hälfte der Differenz zwischen zwei benachbarten Schwellwertspannungspegel zu ändern.
8. Verfahren nach Anspruch 6, dadurch gekennzeichnet,
- daß gemäß dem Schritt des Löschsens der Zelle die Zelle mit einem Löschimpuls abwechselnd gepulst und der Strom durch die Zelle gelesen wird, daß mit Pulsen und Lesen fortgefahren wird, bis der Strompegel den Pegel entsprechend dem gewünschten Basisschwellwertpegel erreicht, wobei die Größe und die Dauer jedes Löschimpulses so gewählt wird, daß der erste Löschimpuls ungenügend ist, die Zelle voll zu löschen, und daß jeder nachfolgende Löschimpuls um ein festgelegtes Maß hinsichtlich seiner Stärke vergrößert wird, bis die Zelle vollständig gelöscht ist.
9. Verfahren nach Anspruch 7, dadurch gekennzeichnet,
- daß gemäß dem Schritt des Löschsens der Zelle die Zelle mit einem Löschimpuls abwechselnd gepulst und ihr effektiver Schwellwertpegel gelesen wird, daß mit dem Pulsen und Lesen so lange fortgefahren wird, bis entweder der effektive Schwellwertpegel den gewünschten Basisschwellwertpegel erreicht oder die Anzahl der angelegten Löschimpulse eine voreingestellte Anzahl überstiegen hat.
10. Verfahren nach Anspruch 5, dadurch gekennzeichnet,
- daß die Zellen der Anordnung Blöcke von zu adressierenden Zellen bilden und
- daß der zusätzliche Schritt der Speicherung des Zählstandes (S) in Zählstandsspeicherzellen innerhalb jedes Blocks der Zellen erfolgt, zu denen der Zählstand gehört.
11. Verfahren nach Anspruch 10, dadurch gekennzeichnet,
- daß der Zählstand (S) zeitweise in einer Regi-

sterdatei gespeichert wird,
daß anschließend der adressierte Block der
Zellen gelöscht und der Zählstand (S) um Eins
erhöht wird, woraufhin der erhöhte Zählstand
erneut in die Zählspeicherzellen des adressier- 5
ten Blockes geschrieben werden.

12. Verfahren nach Anspruch 6 und 10

mit folgenden Schritten der Löschung eines adres- 10
sierten Blockes:

die Zellen werden während einer vorbestimm-
ten Zeit und mit genügender Spannung gepulst,
um die Schwellwertspannung zu ändern, aber
ungenügend, um die Zellen vollständig zu 15
löschen,
danach wird der Strom durch eine vorgewählte
Anzahl (m+n) von Zellen gelesen, um ihre geän-
derte Schwellwertspannung festzustellen,
die Puls- und Leseschritte werden mehrfach 20
wiederholt, wobei bei jeder Wiederholung des
Impulsschrittes die vorbestimmte Spannung um
einen Wert erhöht wird, der oberhalb des letzten
Pulsschrittes ist,
die Puls- und Leseschritte werden bei dem 25
ersten Vorkommen einer der nachfolgenden
Bedingungen beendet:
die effektive Schwellwertspannung jeder der
vorgewählten Anzahl (m+n) der Zellen hat den
Basispegel erreicht, oder 30
der Pulsschritt ist mit einer vorgewählten maxi-
malen Anzahl wiederholt worden, oder
eine vorbestimmte maximale Spannung für
einen Impuls ist bei dem letztvergangenen Puls-
schritt erreicht worden, oder 35
eine Anzahl (N) von Zellen der vorgewählten
Anzahl (m+n) der Zellen, die nicht voll gelöscht
verbleiben, ist gleich oder kleiner einer zulässi-
gen Anzahl von nicht gelöschten Zellen.

13. Verfahren nach Anspruch 12,

dadurch gekennzeichnet, daß die vorgewählte
Anzahl (m+n) der Zellen bedeutend kleiner als die
Gesamtanzahl (mxn) der Zellen des adressierten
Blocks ist. 45

14. Verfahren nach Anspruch 12 oder 13,

dadurch gekennzeichnet, daß der Zählstand (S)
gleich der Gesamtanzahl von Löschkzyklen, den die
Zellen des adressierten Blocks durchgemacht 50
haben, akkumuliert und in Zählstandspeicherzellen
innerhalb jedes Blockes der Zellen gespeichert wird,
zu denen der Zählstand (S) gehört.

15. Vorrichtung in der Form einer Anordnung von elek- 55

trisch änderbaren Speicherzellen mit einer Einrich-
tung zur Adressierung individueller Speicherzellen
zum Lesen und Ändern ihrer Zustände, mit folgen-

den Merkmalen:

jede Zelle umfaßt einen Feldeffekttransistor mit
einer Source-Region (13), einer Drain-Region
(15), einer Kanalregion (17), einem schwim-
menden Gate (19) und einem Steuergate (23),
der Feldeffekttransistor besitzt einen Schwell-
wertspannungspegel, der ein gegebener Pegel
bei Abwesenheit einer Nettoladung auf dem
schwimmenden Gate darstellt, aber gemäß
einem Betrag von Nettoladung variabel ist, die
von dem schwimmenden Gate geführt wird,
die Vorrichtung umfaßt eine Einrichtung zur
Errichtung einer Mehrzahl von effektiven
Schwellwertspannungspegeln (V_{T1}) oberhalb
von Zwei, die einer Mehrzahl von individuellen
feststellbaren Zuständen ("0", "1", "2", "3") der
Zelle oberhalb von Zwei entspricht,
ferner eine Einrichtung zur Voreinstellung der
effektiven Schwellwertspannung der adres-
sierten Zelle auf einen vorbestimmten Pegel durch
Änderung des Ladungsbetrags auf dem
schwimmenden Gate und
eine Einrichtung zur Einstellung der adres-
sierten Zelle auf einen ihrer Zustände der Mehrzahl
möglicher Zustände durch Änderung des
Ladungsbetrags auf dem schwimmenden Gate,
bis die effektive Schwellwertspannung im
wesentlichen gleich dem einen Pegel der Mehr-
zahl der effektiven Schwellwertspannungspe-
gel ist,
dadurch gekennzeichnet,
daß eine Majorität der Mehrzahl der effektiven
Schwellwertpegel oder mindestens zwei effe-
ktive Schwellwertpegel (2, 3) von einer nettopo-
sitiven Ladung auf dem schwimmenden Gate
(19) resultieren.

16. Vorrichtung nach Anspruch 15,

dadurch gekennzeichnet,

daß der Feldeffekttransistor vom Typ mit
gespaltenem Kanal mit einem ersten Kanalteil
(L_1) und einem zweiten Kanalteil (L_2) ist,
daß das schwimmende Gate (19) benachbart
der Drain-Region (15) angeordnet ist und sich
über den ersten Teil (N_1) der Kanalregion (17)
erstreckt, wobei eine erste dielektrische
Gate-Schicht (21) dazwischen in einer Weise
vorgesehen ist, um den Betrag der Leitung
durch den ersten Kanalteil (L_1) gemäß einem
Pegel der Elektronenladung zu steuern, die auf
dem schwimmenden Gate (19) gespeichert ist,
daß das Steuer-Gate (23) sich über mindestens
einen Teil des schwimmenden Gates (19)
erstreckt und von diesem isoliert (27) ist, wobei
ein erster Transistor (T_1) gebildet wird, der den
Schwellwertspannungspegel (V_{T1}) aufweist,

- der bei Abwesenheit der Nettoladung auf dem schwimmenden Gate (19) als der gegebene Schwellwert betrachtet wird, der jedoch gemäß einem Betrag von Nettoladung variabel ist, die von dem schwimmenden Gate (19) geführt wird, 5
- daß das Steuer-Gate (23) auch benachbart der Source-Region (13) angeordnet ist und sich über den zweiten Teil (L_2) der Kanalregion (17) erstreckt, wobei eine zweite dielektrische Gate-Schicht (25) dazwischen in einer Weise vorgesehen ist, um die Leitung durch den zweiten Kanalteil (L_2) gemäß einem Pegel der Spannung zu steuern, die an dem Steuer-Gate (23) anliegt, wobei ein zweiter Transistor (T_2) mit einer festgelegten Schwellwertspannung (V_{T2}) gebildet wird, und 10
- daß der durch den ersten Transistor (T_1) gegebene Schwellwertpegel (V_{T1}) genügend hoch errichtet wird, so daß mindestens zwei programmierbare Schwellwertpegel von einer nettopositiven Ladung auf dem schwimmenden Gate (19) gebildet werden. 15
17. Vorrichtung gemäß Anspruch 16, 25
dadurch gekennzeichnet,
daß der gegebene Schwellwertpegel (V_{T1}) des ersten Transistors (T_1) größer als der festgelegte Schwellwert (V_{T2}) des zweiten Transistors (T_2) ist. 30
18. Vorrichtung nach den Ansprüchen 16 oder 17, 35
dadurch gekennzeichnet,
daß die gegebenen Schwellwertpegel (V_{T1}) des ersten Transistors (T_1) mindestens 3 V beträgt. 40
19. Vorrichtung nach Anspruch 18,
dadurch gekennzeichnet,
daß der festgelegte Schwellwertpegel (V_{T2}) des zweiten Transistorteils (T_2) ungefähr 1 V beträgt. 45
20. Vorrichtung nach einem der Ansprüche 16 bis 17, 50
dadurch gekennzeichnet,
daß der gegebene Schwellwertpegel (V_{T1}) des ersten Transistors (T_1) genügend hoch errichtet wird, damit die Majorität oder alle programmierbaren Schwellwertpegel von einer nettopositiven Ladung auf dem schwimmenden Gate (19) resultieren. 55
21. Vorrichtung nach einem der Ansprüche 16 bis 20, 50
dadurch gekennzeichnet,
daß die Mehrzahl der programmierbaren Schwellwertpegel des ersten Transistors (T_1) genau Drei beträgt.
22. Vorrichtung nach einem der Ansprüche 16 bis 20, 55
dadurch gekennzeichnet,
daß die Mehrzahl der programmierbaren Schwellwertpegel des ersten Transistors (T_1) genau Vier

beträgt.

23. Vorrichtung nach einem der Ansprüche 15 bis 22, dadurch gekennzeichnet,

daß die Voreinstellung und die Einstelleinrichtung folgendes umfassen:
eine Löscheinrichtung ist elektrisch mit mindestens einer Löschelektrode einer adressierten Zelle verbunden, um die Entfernung von Elektronenladung von dem schwimmenden Gate (19) so lange zu veranlassen, bis das schwimmende Gate (19) positiv bis auf einen maximalen Betrag geladen wird,
eine Programmiereinrichtung ist mit mindestens der Source- (13) und Drain-Region (15) sowie dem Steuer-Gate (23) elektrisch verbunden, um die Zufuhr von Elektronenladung auf das schwimmende Gate (19) zu bewirken, um den Betrag der darauf geladenen Elektronenladung von der maximalen positiven Ladung auf einen gewünschten effektiven Schwellwertspannungspegel zu bringen,
eine Leseeinrichtung ist mit mindestens der Source- (13) und Drain-Region (15) sowie dem Steuer-Gate (23) elektrisch verbunden, um den Pegel der auf dem schwimmenden Gate (19) gespeicherten Ladung zu lesen und zu bestimmen, auf welchem der effektiven Schwellwertspannungspegel er liegt, wobei der individuell feststellbare Zustand ("0", "1", "2", "3") der Zelle gelesen wird.

24. Vorrichtung nach Anspruch 23, dadurch gekennzeichnet,

daß die Löscheinrichtungen gleichzeitig mit Löschelektroden oder Gates von Speicherzellen innerhalb eines Blocks einer Mehrzahl von unterschiedlichen Zellenblöcken verbindbar sind, die in Vielfachreihen und -spalten der Anordnung vorliegen, um gleichzeitig die Elektronenladung von den schwimmenden Gates (19) der Zellen innerhalb eines Blocks zu ändern,
daß die Löscheinrichtung eine Impulsfolge zunehmender Stärke an die Löschelektroden oder -Gates des einen Blockes anlegen, bis mindestens entweder

- (i) die Schwellwertpegel (V_{T1}), welche von der Leseeinrichtung zwischen den Löschimpulsen festgestellt werden, und in mindestens einem voreingestellten Anteil ($N \leq X$) der Zellen des einen Blockes existieren, einen Löschepegel erreichen oder
(ii) eine voreingestellte maximale Anzahl (n_{\max}) von Löschimpulsen vorgekommen

sind, und

daß die Programmierereinrichtung mit einer adressierten Zelle innerhalb der Anordnung der Zellen verbindbar ist, um die Elektronenladung auf ihrem schwimmenden Gate (19) durch Anlage einer Reihenfolge von Programmierimpulsen an die adressierte Zelle zu erhöhen, bis mindestens entweder

- (iii) der Schwellwertpegel, der von der Leseeinrichtung zwischen den Programmierimpulsen festgestellt wird und in der adressierten Zelle existiert, von dem Löschepegel auf im wesentlichen einen Wert angehoben wird, der einen gewünschten Pegel der Mehrzahl der vorbestimmten Schwellwertpegel (V_{T1}) gleichkommt, oder
- (iv) eine voreingestellte Anzahl von Programmierimpulsen angelegt worden ist.

25. Vorrichtung nach Anspruch 24,

dadurch gekennzeichnet, daß die Programmierereinrichtung eine Einrichtung zur Begrenzung der Stärke eines jeden Programmierimpulses der Folge auf einen Wert umfaßt, der kleiner als der Wert ist, welcher den Schwellwertpegel des ersten Transistors (T_1) des Anspruchs 16 um die Hälfte der Differenz zwischen benachbarten effektiven Schwellwertspannungspegeln der Mehrzahl verschiebt.

26. Vorrichtung nach einem der Ansprüche 15 bis 25, dadurch gekennzeichnet,

daß eine Leseeinrichtung elektrisch mit mindestens der Source- (13) und Drain-Region (15) und dem Steuer-Gate (23) verbunden ist, um den Pegel der auf dem schwimmenden Gate (19) gespeicherten Ladung zu lesen, daß die Leseeinrichtung eine Mehrzahl von Bezugsspannungsquellen umfaßt, die jeweils der Mehrzahl der effektiven Schwellwertspannungspegel (V_{T1}) entsprechen, und daß eine Einrichtung zum gleichzeitigen Vergleichen des durch die adressierte Zelle fließenden Stroms mit der jeweiligen Bezugsstromquelle vorgesehen ist, wobei der Speicherzustand ("0", "1", "2", "3") der adressierten Zelle bestimmt wird.

Revendications

1. Procédé de modification d'un état de mémoire d'une cellule adressée d'une matrice, la matrice de cellules de mémoire modifiable électriquement comportant un moyen pour adresser des cellules de

mémoire individuelles pour lire et modifier leurs états, chaque cellule comportant un transistor à effet de champ avec une grille flottante (19) et ayant un niveau de tension de seuil qui est un niveau donné en l'absence de charge nette sur ladite grille flottante, mais qui est variable en fonction de la charge électrique nette portée par ladite grille flottante ; comprenant les étapes :

d'établissement d'une pluralité, supérieure à deux, de niveaux de tension de seuil efficaces (V_{T1}) qui correspondent à une pluralité, supérieure à deux, d'états individuellement détectables ("0", "1", "2", "3") de la cellule ; de préréglage à un niveau initial prédéterminé, de la tension de seuil efficace de la cellule adressée, en modifiant la charge électrique sur la grille flottante ; et de mise de la cellule adressée dans l'un de sa dite pluralité d'états par modification de la charge électrique sur la grille flottante jusqu'à ce que sa tension de seuil efficace soit sensiblement égale à l'un de ladite pluralité de niveaux de tension de seuil efficace ; caractérisé en ce qu'une majorité de ladite pluralité de niveaux de seuil efficace, ou au moins deux niveaux de seuil efficace ("2", "3") résultent d'une charge positive nette sur la grille flottante.

2. Procédé selon la revendication 1,

dans lequel l'étape d'établissement d'une pluralité de niveaux de tension de seuil efficace (V_{T1}) comprend l'établissement d'au moins quatre de ces niveaux de tension de seuil efficace.

3. Procédé selon la revendication 1 ou 2,

dans lequel l'étape d'établissement de niveaux de tension de seuil comprend la sélection de la totalité de ladite pluralité de niveaux de tension de seuil efficace pour qu'ils résultent d'une charge nette positive sur la grille flottante.

4. Procédé selon l'une quelconque des revendications 1 à 3,

dans lequel le niveau de seuil donné des transistors de cellule de mémoire est d'au moins trois volts.

5. Procédé selon l'une quelconque des revendications 1 à 4,

qui comprend l'étape additionnelle de cumul d'un compte (S) égal à un nombre total de fois où la cellule a été prérégulée.

6. Procédé selon l'une quelconque des revendications 1 à 5, dans lequel l'étape de préréglage est une étape d'effacement effectuée en supprimant la

charge négative de la grille flottante (19) de manière à abaisser la tension de seuil efficace de la cellule à un niveau de base qui est plus faible que le niveau le plus faible correspondant à ladite pluralité d'états de mémoire détectables de la cellule ; et

dans lequel l'étape de mise dans un état est une étape de programmation effectuée en ajoutant une charge négative à la grille flottante (19).

7. Procédé selon la revendication 6, dans lequel l'étape de programmation de la cellule à l'un quelconque de sa dite pluralité de niveaux de tension de seuil comprend l'application alternée à la cellule d'une brève impulsion de programmation et ensuite la lecture du courant dans la cellule, la poursuite de l'application d'impulsions et de la lecture jusqu'à ce que le niveau de courant atteigne celui correspondant à l'un souhaité de ladite pluralité de niveaux de seuil efficace, chacune desdites brèves impulsions de programmation étant insuffisante pour modifier ladite tension de seuil efficace de plus qu'environ la moitié de la différence entre deux quelconques, adjacents, de ladite pluralité de niveaux de tension de seuil.

8. Procédé selon la revendication 6, dans lequel l'étape d'effacement de la cellule comprend l'application alternée à la cellule d'une impulsion d'effacement et ensuite la lecture du courant dans la cellule, la poursuite de l'application d'impulsions et de la lecture jusqu'à ce que le niveau de courant atteigne celui correspondant au niveau de seuil de base, l'amplitude et la durée de chacune de ces impulsions d'effacement étant choisie pour que la première impulsion d'effacement soit insuffisante pour effacer complètement ladite cellule et pour que chaque impulsion d'effacement ultérieure ait son amplitude accrue d'un incrément fixe jusqu'à ce que la cellule soit complètement effacée.

9. Procédé selon la revendication 7, dans lequel l'étape d'effacement de la cellule comprend l'application alternée à la cellule d'une impulsion d'effacement et la lecture de son niveau de seuil efficace, la poursuite de l'application d'impulsions et de la lecture jusqu'à ce que, soit ledit niveau de seuil efficace atteigne le niveau de seuil de base voulu, soit le nombre d'impulsions d'effacement appliqué dépasse un nombre prédéterminé.

10. Procédé selon la revendication 5,

dans lequel lesdites cellules de ladite matrice forment des blocs de cellules destinés à être adressés ; et

comprenant en outre l'étape additionnelle de mémorisation dudit compte (S) dans des cellules de mémorisation de compte à l'intérieur de

chaque bloc de cellules auquel le compte se rapporte.

11. Procédé selon la revendication 10, dans lequel ledit compte (S) est mémorisé temporairement dans un fichier de registres, puis le bloc de cellules adressé est effacé et le compte (S) est incrémenté de un, après quoi le compte incrémenté est réécrit dans lesdites cellules de mémorisation de compte du bloc adressé.

12. Procédé selon les revendications 6 et 10,

dans lequel l'étape d'effacement d'un bloc adressé comprend :

l'application d'impulsion aux cellules pendant un temps prédéterminé et avec une tension suffisante pour modifier la tension de seuil mais insuffisante pour effacer complètement lesdites cellules ;

puis la lecture du courant dans un nombre sélectionné ($m + n$) de cellules pour vérifier leur tension de seuil modifiée ; et

la répétition des étapes d'application d'impulsion et de lecture plusieurs fois, chaque répétition de l'étape d'application d'impulsion augmentant la tension prédéterminée d'un incrément au-dessus de celle de la dernière étape d'application d'impulsion ;

l'arrêt des étapes d'application d'impulsion et de lecture lors de la première occurrence de l'une quelconque des conditions suivantes :

la tension de seuil efficace de chacune dudit nombre sélectionné ($m + n$) de cellules a atteint le niveau de base ; ou l'étape d'application d'impulsion a été répétée un nombre maximal prédéterminé de fois, ou bien ;

une tension maximale prédéterminée pour une impulsion a été atteinte dans l'étape d'application d'impulsion la plus récente ; ou bien

un nombre (N) de cellules dudit nombre sélectionné ($m + n$) de cellules qui demeurent non complètement effacées est égal ou inférieur à un nombre acceptable de cellules non effacées.

13. Procédé selon la revendication 12,

dans lequel le nombre sélectionné ($m + n$) de cellules est plus petit d'une manière significative, que le nombre total ($m \times n$) de cellules du bloc adressé.

14. Procédé selon la revendication 12 ou 13,

dans lequel le compte (S), égal au nombre total de cycles d'effacement que les cellules du bloc adressé ont subi, est cumulé et mémorisé dans des cellules de mémorisation de compte à l'intérieur de chaque bloc de cellules auquel le compte (S) se rapporte.

15. Dispositif sous forme de matrice de cellules de mémoire modifiables électriquement comportant un moyen d'adressage pour adresser les cellules de mémoire individuelles pour lire et modifier leurs états, chaque cellule comprenant un transistor à effet de champ avec :

une région de source (13) ;
 une région de drain (15) ;
 une région de canal (17) ;
 une grille flottante (19) ; et
 une grille de commande (23) ;
 le transistor à effet de champ ayant un niveau de tension de seuil qui est un niveau donné en l'absence de charge nette sur ladite grille flottante mais qui est variable en fonction de la valeur de la charge nette portée par ladite grille flottante ;
 le dispositif comprenant :
 un moyen pour établir une pluralité, supérieure à deux, de niveaux de tension de seuil efficaces (V_{T1}) qui correspondent à une pluralité, supérieure à deux, d'états individuellement détectables ("0", "1", "2", "3") de la cellule ;
 un moyen pour mettre à un niveau initial prédéterminé, la tension de seuil efficace de la cellule adressée, en modifiant la charge électrique sur la grille flottante ; et
 un moyen pour mettre la cellule adressée dans l'un de sa dite pluralité d'états par modification de la charge électrique sur la grille flottante jusqu'à ce que sa tension de seuil efficace soit sensiblement égale à l'un de ladite pluralité de niveaux de tension de seuil efficace ;
 caractérisé en ce qu'une majorité de ladite pluralité de niveaux de seuil efficace, ou au moins deux niveaux de seuil efficace ("2", "3") résultent d'une charge positive nette sur la grille flottante (19).

16. Dispositif selon la revendication 15,

dans lequel ledit transistor à effet de champ est du type à canal en deux parties avec une première partie de canal (L_1) et une seconde partie de canal (L_2) ; et
 dans lequel ladite grille flottante (19) est située adjacente à ladite région de drain (15) et s'étend dans ladite première partie (L_1) de ladite région de canal (17), avec une première couche diélectrique de grille (21) entre elles, de manière à commander la valeur de la conduction dans la première partie de canal (L_1) en fonction d'un niveau de charge d'électrons emmagasinés sur ladite grille flottante (19) ;
 dans lequel ladite grille de commande (23) s'étend sur au moins une partie de la grille flottante (19) et en est isolée (27), en formant ainsi

un premier transistor (T_1) ayant ledit niveau de tension de seuil (V_{T1}) qui est donné en l'absence de charge nette sur ladite grille flottante (19) mais qui est variable en fonction d'une charge électrique nette portée par ladite grille flottante (19) ;
 dans lequel ladite grille de commande (23) est également située adjacente à ladite région de source (13) et s'étend dans ladite seconde partie (L_2) de ladite région de canal (17), avec une seconde couche diélectrique de grille (25) entre elles, de manière à commander la conduction dans la seconde partie de canal (L_2) en fonction d'un niveau de tension appliqué à la grille de commande (23), en formant ainsi un second transistor (T_2) ayant un niveau de seuil fixe (V_{T2}) ; et
 dans lequel ledit niveau de seuil (V_{T1}) donné par ledit premier transistor (T_1) est établi suffisamment haut pour qu'au moins deux niveaux de seuil programmables résultent d'une charge positive nette sur la grille flottante (19).

17. Dispositif selon la revendication 16,
 dans lequel ledit niveau de seuil donné (V_{T1}) dudit premier transistor (T_1) est plus grand que le seuil fixe (V_{T2}) dudit second transistor (T_2).

18. Dispositif selon la revendication 16 ou 17,
 dans lequel ledit niveau de seuil donné (V_{T1}) dudit premier transistor (T_1) est d'au moins 3 volts.

19. Dispositif selon la revendication 18,
 dans lequel le niveau de seuil fixe (V_{T2}) de la seconde partie de transistor (T_2) est d'environ 1 volt.

20. Dispositif selon l'une quelconque des revendications 16 à 17,
 dans lequel ledit niveau de seuil donné (V_{T1}) dudit premier transistor (T_1) est établi suffisamment haut pour que la majorité, ou la totalité, desdits niveaux de seuil programmables résultent d'une charge positive nette sur la grille flottante (19).

21. Dispositif selon l'une quelconque des revendications 16 à 20,
 dans lequel ladite pluralité de niveaux de seuil programmable du premier transistor (T_1) est exactement de trois.

22. Dispositif selon l'une quelconque des revendications 16 à 20,
 dans lequel ladite pluralité de niveaux de seuil programmable du premier transistor (T_1) est exactement de quatre.

23. Dispositif selon l'une quelconque des revendications 15 à 22,

dans lequel ledit moyen de mise à un niveau initial et de mise dans un état comprend :

un moyen d'effacement connecté électriquement à au moins une électrode d'effacement d'une cellule adressée pour provoquer l'élimination de la charge d'électrons de la grille flottante (19) jusqu'à ce que la grille flottante (19) devienne chargée positivement jusqu'à une valeur maximale ;

un moyen de programmation connecté électriquement à au moins ladite source (13), ledit drain (15) et ladite grille de commande (23) pour provoquer l'addition d'une charge d'électrons à la grille flottante (19) pour amener la valeur de la charge d'électrons qui y est emmagasinée, de ladite charge positive maximale à l'un souhaité desdits niveaux de tension de seuil efficace ; et

un moyen de lecture connecté électriquement au moins à ladite source (13), audit drain (15) et à ladite grille de commande (23) pour lire le niveau de charge emmagasinée sur la grille flottante (19) et pour déterminer dans lequel desdits niveaux de tension de seuil efficace elle se trouve, pour lire ainsi les états individuellement détectables ("0", "1", "2", "3") de la cellule.

24. Dispositif selon la revendication 23,

dans lequel ledit moyen d'effacement peut être connecté simultanément aux électrodes d'effacement ou aux grilles des cellules de mémorisation à l'intérieur de l'un de plusieurs blocs de cellules distincts dans les multiples rangées et colonnes de la matrice pour modifier simultanément la charge d'électrons des grilles flottantes (19) des cellules dans ledit un bloc ;

dans lequel ledit moyen d'effacement applique une séquence d'impulsions d'amplitudes croissantes auxdites électrodes d'effacement ou auxdites grilles dudit un bloc jusqu'à ce qu'au moins, soit :

(i) les niveaux de seuil (V_{T1}), détectés par ledit moyen de lecture entre lesdites impulsions d'effacement, existant dans au moins une proportion prédéterminée ($N \leq X$) des cellules dudit un bloc, atteint un niveau effacé ; soit

(ii) un nombre maximal prédéterminé (n_{max}) d'impulsions d'effacement a été produit ; et

dans lequel ledit moyen de programmation peut être connecté à une cellule adressée dans la matrice de cellules pour augmenter la charge d'électrons sur sa grille flottante (19) par application d'une séquence d'impulsions de pro-

grammation à la cellule adressée jusqu'à ce qu'au moins, soit :

(iii) le niveau de seuil détecté par ledit moyen de lecture entre lesdites impulsions de programmation existant dans la cellule adressée est élevé par rapport audit niveau d'effacement jusqu'à égaler sensiblement l'un souhaité de ladite pluralité de niveaux de seuil prédéterminés (V_{T1}) ; soit ;
(iv) un nombre maximal prédéterminé d'impulsions de programmation a été appliqué.

25. Dispositif selon la revendication 24,

dans lequel ledit moyen de programmation comprend un moyen pour limiter l'amplitude de chacune de ladite séquence d'impulsions de programmation à une valeur qui est plus petite que celle qui décale le niveau de seuil du premier transistor (T_1) de la revendication 16 de la moitié de la différence entre ceux qui sont adjacents de ladite pluralité de niveaux de tension de seuil efficace.

26. Dispositif selon l'une quelconque des revendications 15 à 25,

comprenant en outre un moyen de lecture connecté électriquement au moins à ladite source (13), audit drain (15) et à ladite grille de commande (23) pour lire le niveau de charge emmagasinée sur la grille flottante (19) ;

dans lequel ledit moyen de lecture comprend : une pluralité de sources de courant de référence correspondant, respectivement, à ladite pluralité de niveaux de tension de seuil efficace (V_{T1}) ; et

un moyen pour comparer simultanément le courant s'écoulant dans la cellule adressée à chacune des sources de courant de référence, pour déterminer ainsi l'état de mémoire ("0", "1", "2", "3") de la cellule adressée.

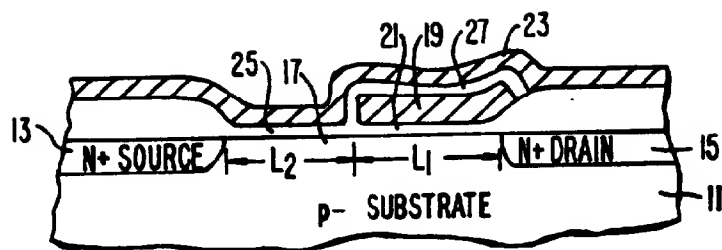


FIG. 1.

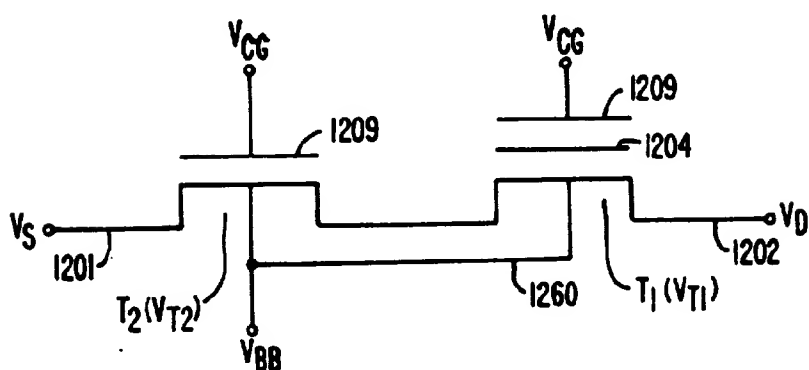


FIG. 2a.

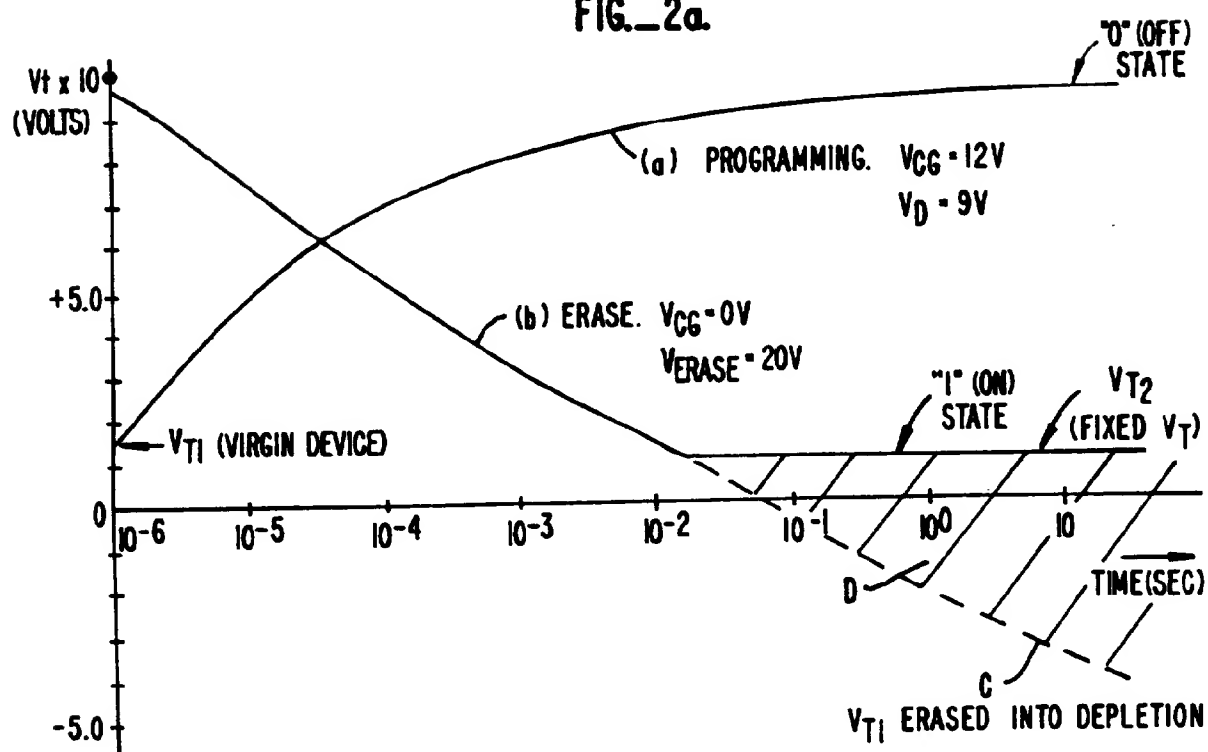


FIG. 2b.

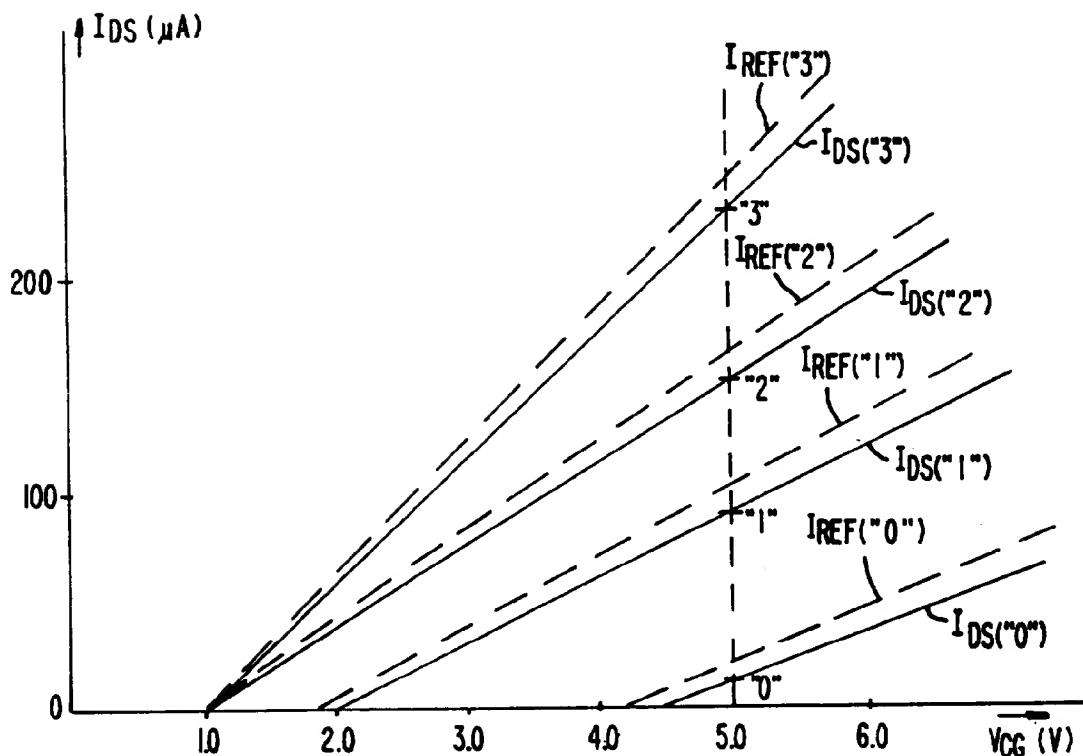


FIG. 2c.

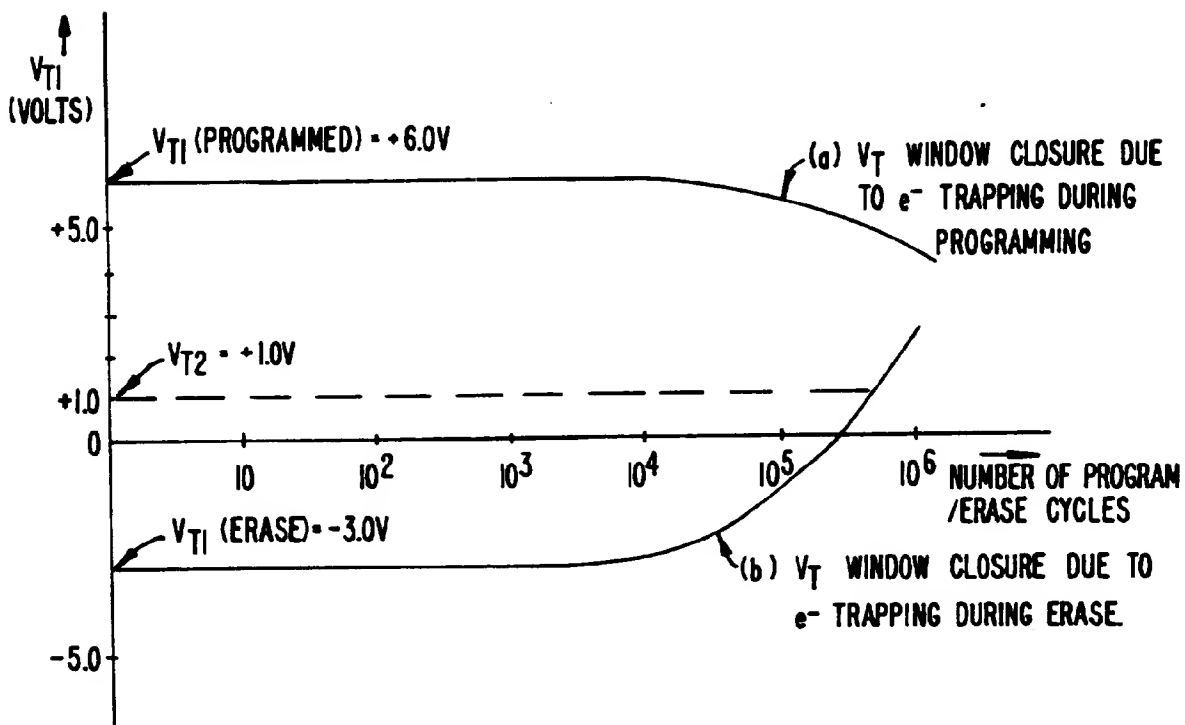


FIG. 2d.

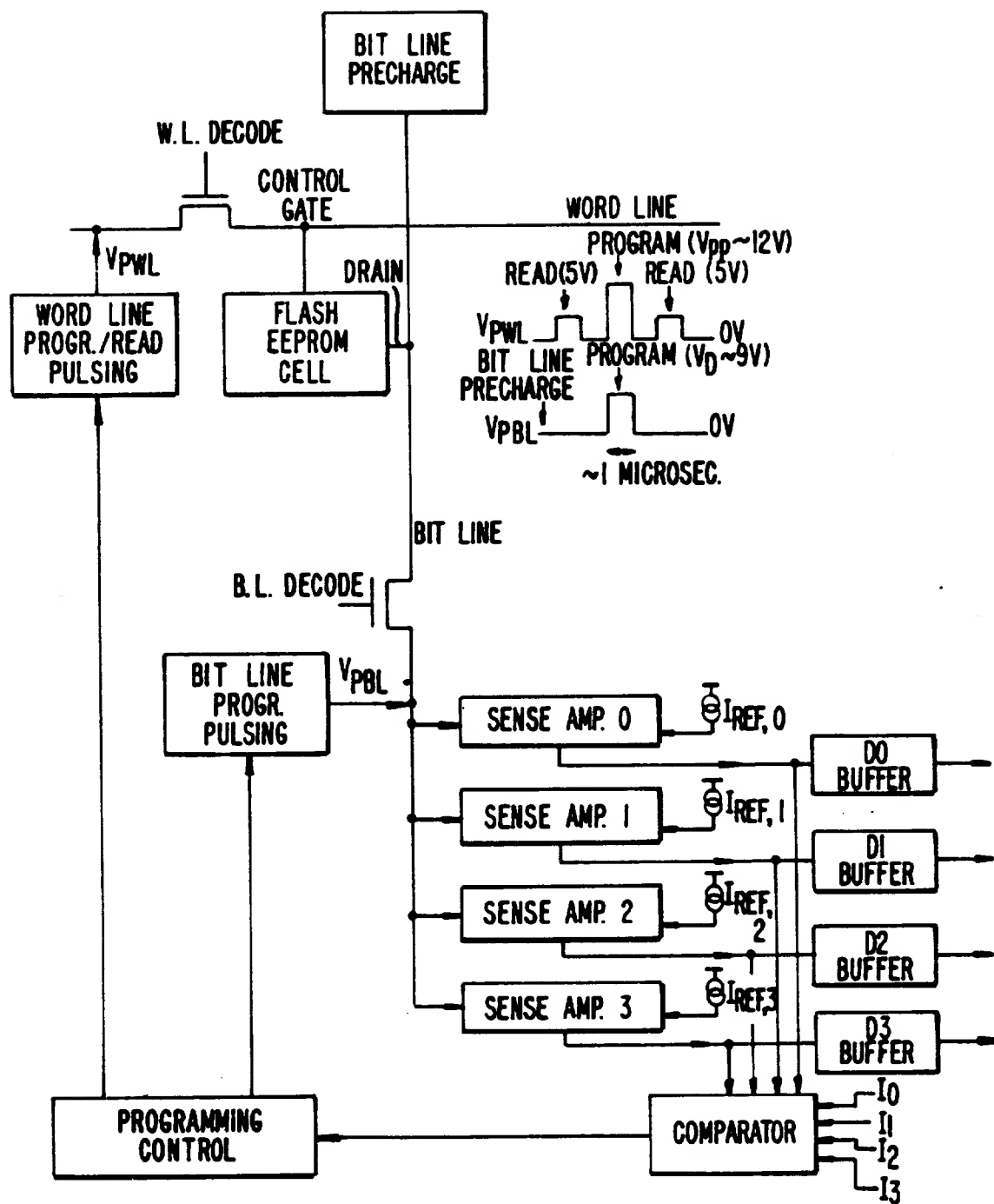


FIG. 2e.

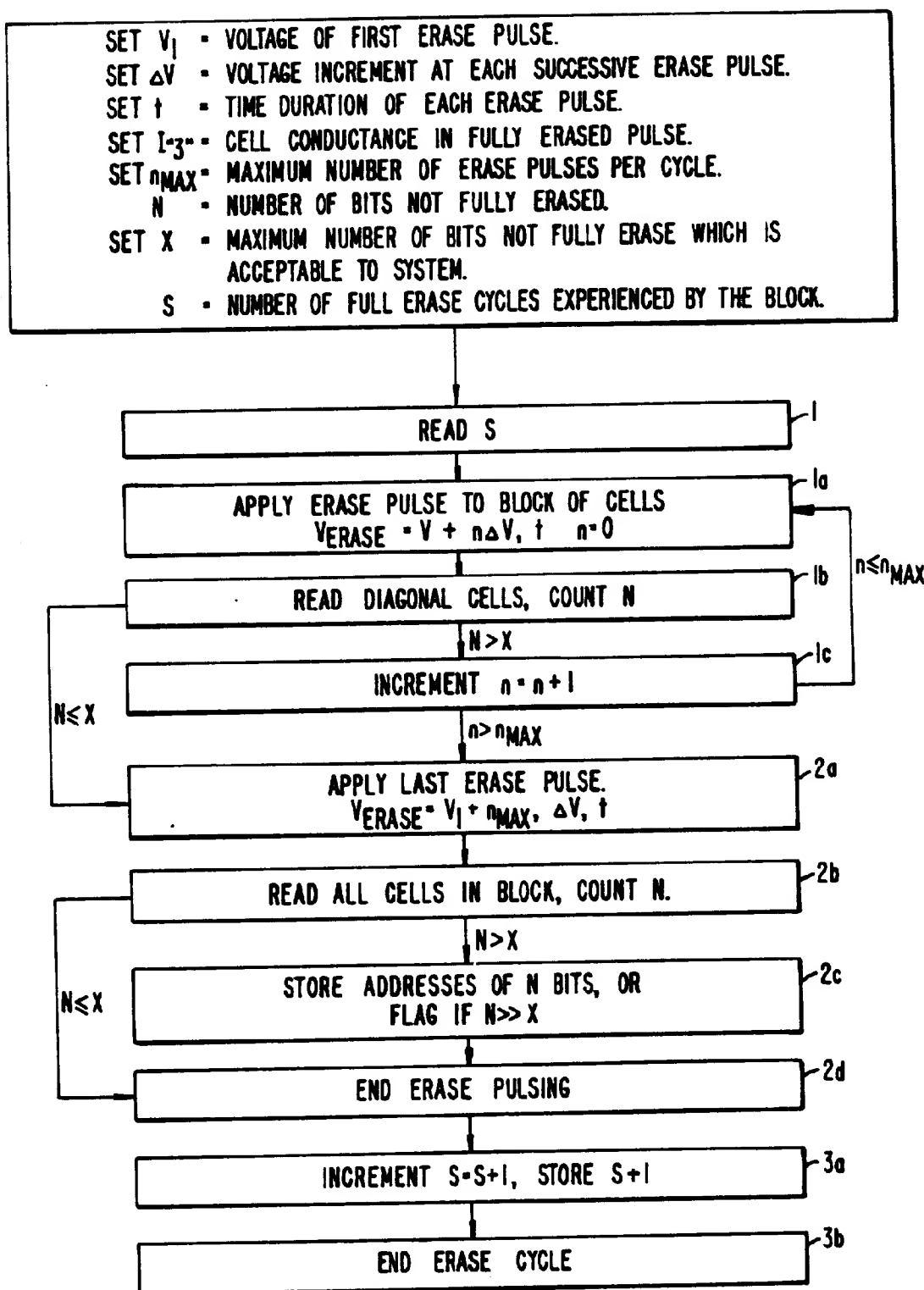


FIG. 3.

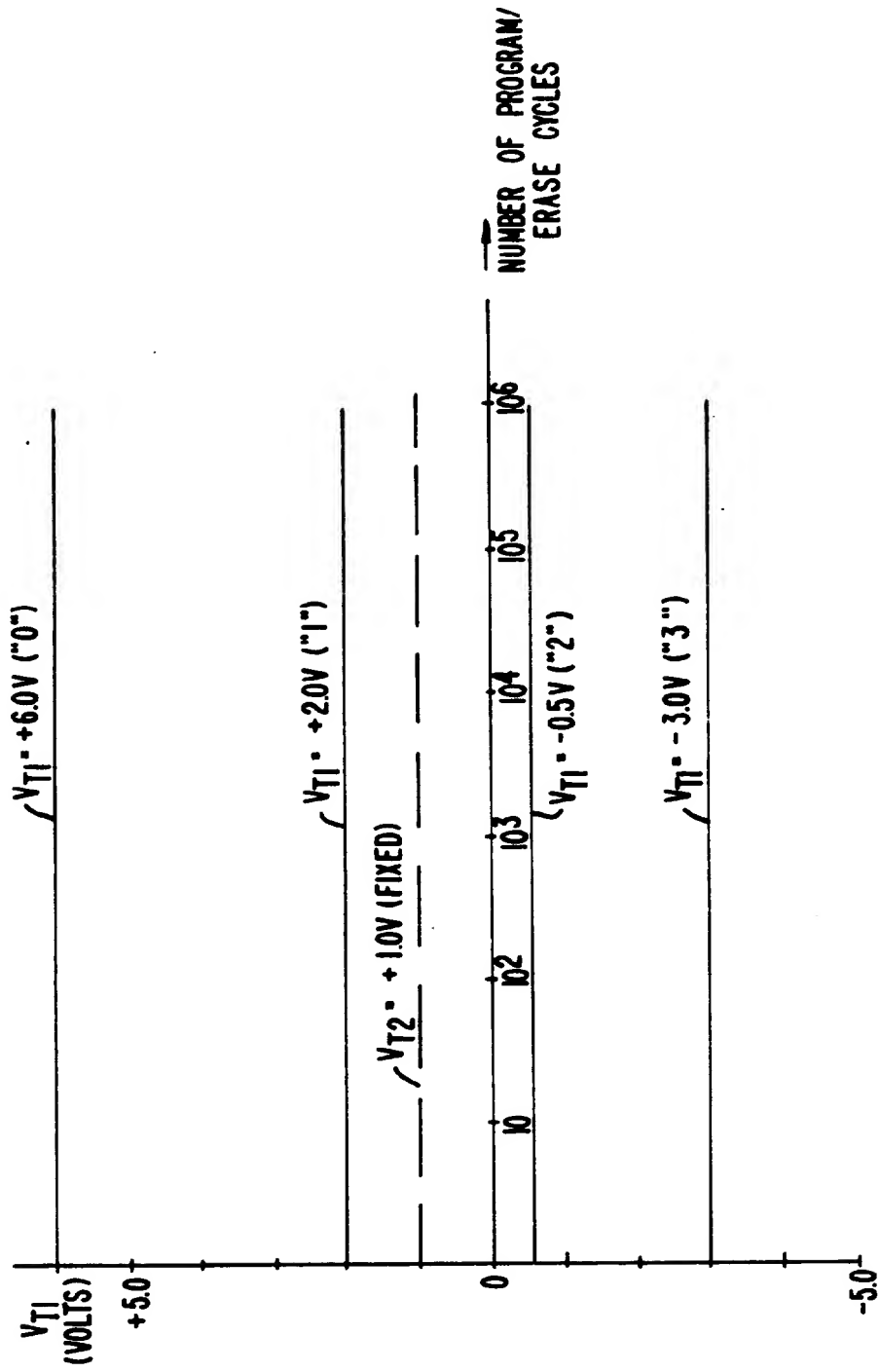


FIG. 4.